THE IMPACT OF NOISE AT DIFFERENT DATA ATTRIBUTES

Nittaya Kerdprasop, Kittisak Kerdprasop, Laksamee Khomnotai
and Thammasak Thianniwet

Data Engineering and Knowledge Discovery (DEKD) Research Unit
School of Computer Engineering Suranaree University of Technology
Nakhon Ratchasima 30000, Thailand
E-mail: nittaya, kerdpras@sut.ac.th

Abstract

Real-world data often suffer from corruptions or noise. The most serious negative impact of noise is that it can reduce machine learning performance in terms of learning accuracy. Most learning algorithms have integrated various approaches to handle noisy data. However, rare research has been conducted to systematically explore the impact of noise, especially when noise occurs at different attributes. We investigate the effect of class noise, noise in principal attributes, and noise in irrelevant attributes to the learning accuracy. Our conclusions can be served as a preliminary step toward the designing of handling mechanisms for a specific kind of noise.

Published in: Proceedings of 1st KMITL International Conference on Integration of Science and Technology for Sustainable Development, Bangkok, Thailand, August 25-26, 2004, pp.275-278.